CLAIMS:

1.   A method of segmentation of a text (512) into text sections and assigning
a topic to each text section on the basis of annotated training data, the method
comprising the steps of:
-    segmenting the text (512) into text sections by making use of statistical models
5   (514) extracted from training data,
-    assigning a topic being indicative of the content of the text section to each text
      section by making use of the statistical models extracted from the training data,
-    generating a structured text by inserting a label as a section heading into the text
      in order to assign the label to the text section,
10   -   providing the structured text to a user (506),
-    processing of modifications of the structured text in response to a user's review.

2.   The method according to claim 1, wherein the topic assigned to a text
section is further assigned to a set of labels (410), one of which being assigned to the
15   text section and inserted as section heading into the text.

3.   The method according to claim 1 or 2, wherein providing the structured
text to a user further comprises for each text section providing the set of labels (410)
assigned to the topic that is assigned to the text section.

20

4.   The method according to any one of the claims 1 to 3, wherein the text
modification comprises a modification of the segmentation of the text into sections
and/or a modification of the assignment between a label and a text section.

5.   The method according to claim 3 or 4, wherein the modification of the structured text comprises:

- assigning a label to a text section by selecting one label (412, 414,...) of the set
5   of labels (410) assigned to the topic that is assigned to the text section,

- re-defining a section boundary by selecting an assigned label (406) at a first position in the text and moving the assigned label to a second position within the text, the second position defining the section boundary, and the selected label defining the section heading,

10   - entering a label and assigning the entered label to the text section.

6.   The method according to any one of the claims 1 to 5, wherein the processing of modifications of the structured text (518) comprises performing modifications in the text in response to the user's review and successively triggering the

15   steps of:

- re-segmenting the text into text sections by making use of the statistical models (514) extracted from the training data and by making reference to the performed modifications,

- re-generating a structured text (518) by inserting a label as a section heading
20   into the text by making reference to the performed modifications, assigning the label to the text section and providing the structured text to the user for review.

7.   The method according to any one of the claims 1 to 6, wherein the processing of modifications of the structured text comprises replacing a text portion by

25   a label within the text, when the replaced text portion is identified as a formulation describing a section heading.

8.   The method according to any one of the claims 1 to 7, wherein the granularity of the text segmentation is controlled by the user by means of a

30   customizable granularity parameter.

9.   The method according to any one of the claims 1 to 8, wherein the modifications of the structured text are logged and analyzed in order to adapt the statistical models.

10.   A text segmentation system (500) for segmenting a text (512) into text sections and assigning a topic to each text section on the basis of annotated training data, the text segmentation system comprising:

- means for segmenting the text (512) into text sections by making use of statistical models (514) extracted from the training data,

- means for assigning a topic being indicative of the content of the text section to each text section by making use of the statistical models extracted from the training data, the topic being further assigned to a set of labels (410),

- means for generating a structured text (518) by inserting one label of the set of labels as a section heading into the text in order to assign the label to the text section,

- means for providing (520) the structured text to a user (506),

- means for processing (516) of modifications of the structured text in response to a user's review.

11.   The text segmentation system according to claim 10, wherein means for processing of modifications (516) of the structured text (518) are adapted to perform a modification of the segmentation of the text into sections and/or a modification of the assignment between a label and a text section.

12.   The text segmentation system according to claim 10 or 11, wherein means for processing of modifications of the structured text are further adapted to perform:

- assigning a label to a text section by selecting one label (412, 414,...) of the set of labels (410) assigned to the topic that is assigned to the text section,

       - re-defining a section boundary by selecting an assigned label (406) at a first position in the text and moving the assigned label to a second position within the text, the second position defining the section boundary, and the selected label defining the section heading,

5   - entering a label and assigning the entered label to the text section.

13.   The text segmentation system according to any one of the claims 10 to 12, wherein the means for processing of modifications (516) of the structured text (518) are adapted to perform modifications in the text in response to the user's review and

10   further comprising means for successively triggering the steps of:

      - re-segmenting the text into text sections by making use of the statistical models (514) extracted from the training data and by making reference to the performed modifications,

      - re-generating a structured text by inserting a label as a section heading into the

15   text by making reference to the performed modifications, assigning the label to the text section and providing the structured text to a user for review.

14.   The text segmentation system according to any one of the claims 10 to 13, further comprising means for logging and analyzing the performed modifications of

20   the structured text, the means for logging and analyzing being adapted to adapt the statistical models (514).

15.   A computer program product for segmenting a text (512) into text sections and assigning a topic to each text section on the basis of annotated training

25   data, the computer program product comprising program means for:

      - segmenting the text into text sections by making use of statistical models (514) extracted from the training data,

      - assigning a topic being indicative of the content of the text section to each text section by making use of the statistical models extracted from the training data,

30   the topic being further assigned to a set of labels (410),

- generating a structured text (518) by inserting one label (412, 414,...) of the set of labels (410) as a section heading into the text in order to assign the label to the text section,

- providing the structured text to a user (506),

5 - processing of modifications of the structured text (518) in response to a user's review.

16. The computer program product according to claim 15, wherein the program means for processing of modifications of the structured text are adapted to

10 perform a modification of the segmentation of the text into sections and/or a modification of the assignment between a label and a text section, for the modification of the assignment between a label and a text section the program means are further adapted to perform the steps of:

- assigning a label to a text section by selecting one label of the set of labels

15 assigned to the topic that is assigned to the text section,

- re-defining a section boundary by selecting an assigned label at a first position in the text and moving the assigned label to a second position within the text, the second position defining the section boundary, and the selected label defining the section heading,

20 - entering a label and assigning the entered label to the text section.

17. The computer program product according to claim 15 or 16, wherein the program means for processing of modifications of the structured text are adapted to perform modifications in the text in response to the user's review and further

25 comprising program means for successively triggering the steps of:

- re-segmenting the text into text sections by making use of the statistical models extracted from the training data and by making reference to the performed modifications,

- re-generating a structured text (518) by inserting a label as a section heading

30 into the text by making reference to the performed modifications, assigning the

label to the text section and providing the structured text to the user for review.

18.   A user interface (400) for a text segmentation system for segmenting a
text into text sections and assigning a topic to each text section on the basis of
5   annotated training data, the user interface comprising:
  -   means for providing the structured text to a user that has been structured by
      making use of statistical models extracted from the training data,
  -   means for providing a set of labels (410) to the user, the set of labels being
      assigned to each topic that is assigned to each text section,
10 -   input means (408) for processing modifications of the structured text in response
      to a user's review,
  -   means for logging and analyzing processed modifications of the structured text
      in order to train the statistical models.

15 19. .   The user interface according to claim 18, wherein the structured text is
provided to a user by means of a graphical user interface (400; 502) and wherein the
input means (408, 418; 504) are adapted to process modifications of the structured text
in form of the user selecting one label (412, 414,...) of the provided set of labels (410),
the selected label being assigned to the text section.
20
20.   The user interface according to claim 18 or 19, further comprising means
for providing text that has been re-segmented and re-labelled in response to the user's
review by making use of the statistical models and by making reference to the
processed modifications.
25